

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/039,177

1646

DATE: 11/06/98  
TIME: 17:12:47

INPUT SET: S29620.raw

This Raw Listing contains the General  
Information Section and up to the first 5 pages

ENTERED

## SEQUENCE LISTING

1  
2  
3 (1) General Information:  
4  
5 (i) APPLICANT: Kohei MIYAZONO; Takeshe IMAMURA; Peter DEN DIJKE  
6  
7 (ii) TITLE OF INVENTION: ISOLATED ALK-1 PROTEIN, NUCLEIC ACIDS ENCODING  
8 IT, AND USES THEREOF  
9  
10 (iii) NUMBER OF SEQUENCES: 29  
11  
12 (iv) CORRESPONDENCE ADDRESS:  
13 (A) ADDRESSEE: Fulbright & Jaworski L.L.P.  
14 (B) STREET: 805 Third Avenue  
15 (C) CITY: New York City  
16 (D) STATE: New York  
17 (E) COUNTRY: USA  
18 (F) ZIP: 10022  
19  
20 (v) COMPUTER READABLE FORM:  
21 (A) MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb  
22 (B) COMPUTER: IBM PS/2  
23 (C) OPERATING SYSTEM: PC-DOS  
24 (D) SOFTWARE: Wordperfect  
25  
26 (vi) CURRENT APPLICATION DATA:  
27 (A) APPLICATION NUMBER: 09/039,177  
28 (B) FILING DATE: March 13, 1998  
29 (C) CLASSIFICATION: 435  
30  
31 (vii) PRIOR APPLICATION DATA:  
32 (A) APPLICATION NUMBER: PCT/GB93/0236  
33 (B) FILING DATE: November 17, 1993  
34  
35 (vii) PRIOR APPLICATION DATA:  
36 (A) APPLICATION NUMBER: GB 9224057.1  
37 (B) FILING DATE: November 17, 1992  
38  
39 (vii) PRIOR APPLICATION DATA:  
40 (A) APPLICATION NUMBER: GB 9304677.9  
41 (B) FILING DATE: March 8, 1993  
42  
43 (vii) PRIOR APPLICATION DATA:  
44 (A) APPLICATION NUMBER: GB 9304680.3  
45 (B) FILING DATE: March 8, 1993  
46

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/039,177DATE: 11/06/98  
TIME: 17:12:48

INPUT SET: S29620.raw

47 (vii) PRIOR APPLICATION DATA:  
48 (A) APPLICATION NUMBER: 9311047.6  
49 (B) FILING DATE: May 28, 1993  
50  
51 (vii) PRIOR APPLICATION DATA:  
52 (A) APPLICATION NUMBER: 9313763.6  
53 (B) FILING DATE: July 2, 1993  
54  
55 (vii) PRIOR APPLICATION DATA:  
56 (A) APPLICATION NUMBER: 9136099.2  
57 (B) FILING DATE: August 3, 1993  
58  
59 (vii) PRIOR APPLICATION DATA:  
60 (A) APPLICATION NUMBER: 321344.5  
61 (B) FILING DATE: October 15, 1993  
62  
63 (viii) ATTORNEY/AGENT INFORMATION:  
64 (A) NAME: Norman D. Hanson  
65 (B) REGISTRATION NUMBER: 30,946  
66 (C) REFERENCE/DOCKET NUMBER: LUD 5539 - JEL/NDH  
67  
68 (ix) TELECOMMUNICATION INFORMATION:  
69 (A) TELEPHONE: (212) 688-9200  
70 (B) TELEFAX: (212) 838-3884  
71  
72 (2) INFORMATION FOR SEQ ID NO: 1:  
73 (i) SEQUENCE CHARACTERISTICS:  
74 (A) LENGTH: 1984 base pairs  
75 (B) TYPE: nucleic acid  
76 (C) STRANDEDNESS: unknown  
77 (D) TOPOLOGY: linear  
78  
79 (ii) MOLECULE TYPE: cDNA  
80 (iii) HYPOTHETICAL: NO  
81 (iii) ANTI-SENSE: NO  
82 (v) FRAGMENT TYPE: internal  
83 (vi) ORIGINAL SOURCE:  
84 (A) ORGANISM: Homo sapiens  
85 (ix) FEATURE:  
86 (A) NAME/KEY: CDS  
87 (B) LOCATION: 283..1791  
88  
89 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
90  
91 AGGAAACGGT TTATTAGGAG GGAGTGGTGG AGCTGGGCCA GGCAGGAAGA CGCTGGAATA 60  
92  
93 AGAAACATTT TTGCTCCAGC CCCCATCCCA GTCCCGGGAG GCTGCCGCGC CAGCTGCGCC 120  
94  
95 GAGCGAGCCC CTCCCCGGCT CCAGCCCGGT CCGGGGCCGC GCCGGACCCC AGCCCGCCGT 180  
96  
97 CCAGCGCTGG CGGTGCAACT GCGGCCGCGC GGTGGAGGGG AGGTGGCCCC GGTCCGCCGA 240  
98  
99 AGGCTAGCGC CCCGCCACCC GCAGAGCGGG CCCAGAGGGA CC ATG ACC TTG GGC 294

**INPUT SET: S29620.raw**

Met Thr Leu Gly

1

100													Met	Thr	Leu	Gly	
101													1				
102																	
103	TCC	CCC	AGG	AAA	GGC	CTT	CTG	ATG	CTG	CTG	ATG	GCC	TTG	GTG	ACC	CAG	342
104	Ser	Pro	Arg	Lys	Gly	Leu	Leu	Met	Leu	Leu	Met	Ala	Leu	Val	Thr	Gln	
105	5					10					15					20	
106																	
107	GGA	GAC	CCT	GTG	AAG	CCG	TCT	CGG	GGC	CCG	CTG	GTG	ACC	TGC	ACG	TGT	390
108	Gly	Asp	Pro	Val	Lys	Pro	Ser	Arg	Gly	Pro	Leu	Val	Thr	Cys	Thr	Cys	
109					25					30					35		
110																	
111	GAG	AGC	CCA	CAT	TGC	AAG	GGG	CCT	ACC	TGC	CGG	GGG	GCC	TGG	TGC	ACA	438
112	Glu	Ser	Pro	His	Cys	Lys	Gly	Pro	Thr	Cys	Arg	Gly	Ala	Trp	Cys	Thr	
113				40					45					50			
114																	
115	GTA	GTG	CTG	GTG	CGG	GAG	GAG	GGG	AGG	CAC	CCC	CAG	GAA	CAT	CGG	GGC	486
116	Val	Val	Leu	Val	Arg	Glu	Glu	Gly	Arg	His	Pro	Gln	Glu	His	Arg	Gly	
117			55					60					65				
118																	
119	TGC	GGG	AAC	TTG	CAC	AGG	GAG	CTC	TGC	AGG	GGG	CGC	CCC	ACC	GAG	TTC	534
120	Cys	Gly	Asn	Leu	His	Arg	Glu	Leu	Cys	Arg	Gly	Arg	Pro	Thr	Glu	Phe	
121		70					75					80					
122																	
123	GTC	AAC	CAC	TAC	TGC	TGC	GAC	AGC	CAC	CTC	TGC	AAC	CAC	AAC	GTG	TCC	582
124	Val	Asn	His	Tyr	Cys	Cys	Asp	Ser	His	Leu	Cys	Asn	His	Asn	Val	Ser	
125	85					90					95					100	
126																	
127	CTG	GTG	CTG	GAG	GCC	ACC	CAA	CCT	CCT	TCG	GAG	CAG	CCG	GGA	ACA	GAT	630
128	Leu	Val	Leu	Glu	Ala	Thr	Gln	Pro	Pro	Ser	Glu	Gln	Pro	Gly	Thr	Asp	
129					105					110					115		
130																	
131	GGC	CAG	CTG	GCC	CTG	ATC	CTG	GGC	CCC	GTG	CTG	GCC	TTG	CTG	GCC	CTG	678
132	Gly	Gln	Leu	Ala	Leu	Ile	Leu	Gly	Pro	Val	Leu	Ala	Leu	Leu	Ala	Leu	
133				120					125					130			
134																	
135	GTG	GCC	CTG	GGT	GTC	CTG	GGC	CTG	TGG	CAT	GTC	CGA	CGG	AGG	CAG	GAG	726
136	Val	Ala	Leu	Gly	Val	Leu	Gly	Leu	Trp	His	Val	Arg	Arg	Arg	Gln	Glu	
137			135					140					145				
138																	
139	AAG	CAG	CGT	GGC	CTG	CAC	AGC	GAG	CTG	GGA	GAG	TCC	AGT	CTC	ATC		

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/039,177

DATE: 11/06/98  
TIME: 17:12:50

INPUT SET: S29620.raw

	200	205	210	
153				
154				
155	TAT GGC GAA GTG TGG CGG GGC TTG TGG CAC GGT GAG AGT GTG GCC GTC			966
156	Tyr Gly Glu Val Trp Arg Gly Leu Trp His Gly Glu Ser Val Ala Val			
157	215	220	225	
158				
159	AAG ATC TTC TCC TCG AGG GAT GAA CAG TCC TGG TTC CGG GAG ACT GAG			1014
160	Lys Ile Phe Ser Ser Arg Asp Glu Gln Ser Trp Phe Arg Glu Thr Glu			
161	230	235	240	
162				
163	ATC TAT AAC ACA GTA TTG CTC AGA CAC GAC AAC ATC CTA GGC TTC ATC			1062
164	Ile Tyr Asn Thr Val Leu Leu Arg His Asp Asn Ile Leu Gly Phe Ile			
165	245	250	255	260
166				
167	GCC TCA GAC ATG ACC TCC CGC AAC TCG AGC ACG CAG CTG TGG CTC ATC			1110
168	Ala Ser Asp Met Thr Ser Arg Asn Ser Ser Thr Gln Leu Trp Leu Ile			
169	265	270	275	
170				
171	ACG CAC TAC CAC GAG CAC GGC TCC CTC TAC GAC TTT CTG CAG AGA CAG			1158
172	Thr His Tyr His Glu His Gly Ser Leu Tyr Asp Phe Leu Gln Arg Gln			
173	280	285	290	
174				
175	ACG CTG GAG CCC CAT CTG GCT CTG AGG CTA GCT GTG TCC GCG GCA TGC			1206
176	Thr Leu Glu Pro His Leu Ala Leu Arg Leu Ala Val Ser Ala Ala Cys			
177	295	300	305	
178				
179	GGC CTG GCG CAC CTG CAC GTG GAG ATC TTC GGT ACA CAG GGC AAA CCA			1254
180	Gly Leu Ala His Leu His Val Glu Ile Phe Gly Thr Gln Gly Lys Pro			
181	310	315	320	
182				
183	GCC ATT GCC CAC CGC GAC TTC AAG AGC CGC AAT GTG CTG GTC AAG AGC			1302
184	Ala Ile Ala His Arg Asp Phe Lys Ser Arg Asn Val Leu Val Lys Ser			
185	325	330	335	340
186				
187	AAC CTG CAG TGT TGC ATC GCC GAC CTG GGC CTG GCT GTG ATG CAC TCA			1350
188	Asn Leu Gln Cys Cys Ile Ala Asp Leu Gly Leu Ala Val Met His Ser			
189	345	350	355	
190				
191	CAG GGC AGC GAT TAC CTG GAC ATC GGC AAC AAC CCG AGA GTG GGC ACC			1398
192	Gln Gly Ser Asp Tyr Leu Asp Ile Gly Asn Asn Pro Arg Val Gly Thr			
193	360	365	370	
194				
195	AAG CGG TAC ATG GCA CCC GAG GTG CTG GAC GAG CAG ATC CGC ACG GAC			1446
196	Lys Arg Tyr Met Ala Pro Glu Val Leu Asp Glu Gln Ile Arg Thr Asp			
197	375	380	385	
198				
199	TGC TTT GAG TCC TAC AAG TGG ACT GAC ATC TGG GCC TTT GGC CTG GTG			1494
200	Cys Phe Glu Ser Tyr Lys Trp Thr Asp Ile Trp Ala Phe Gly Leu Val			
201	390	395	400	
202				
203	CTG TGG GAG ATT GCC CGC CGG ACC ATC GTG AAT GGC ATC GTG GAG GAC			1542
204	Leu Trp Glu Ile Ala Arg Thr Ile Val Asn Gly Ile Val Glu Asp			
205	405	410	415	420

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/039,177

DATE: 11/06/98  
TIME: 17:12:50

INPUT SET: S29620.raw

```

206
207 TAT AGA CCA CCC TTC TAT GAT GTG GTG CCC AAT GAC CCC AGC TTT GAG 1590
208 Tyr Arg Pro Pro Phe Tyr Asp Val Val Pro Asn Asp Pro Ser Phe Glu
209 425 430 435
210
211 GAC ATG AAG AAG GTG GTG TGT GTG GAT CAG CAG ACC CCC ACC ATC CCT 1638
212 Asp Met Lys Lys Val Val Cys Val Asp Gln Gln Thr Pro Thr Ile Pro
213 440 445 450
214
215 AAC CGG CTG GCT GCA GAC CCG GTC CTC TCA GGC CTA GCT CAG ATG ATG 1686
216 Asn Arg Leu Ala Ala Asp Pro Val Leu Ser Gly Leu Ala Gln Met Met
217 455 460 465
218
219 CGG GAG TGC TGG TAC CCA AAC CCC TCT GCC CGA CTC ACC GCG CTG CGG 1734
220 Arg Glu Cys Trp Tyr Pro Asn Pro Ser Ala Arg Leu Thr Ala Leu Arg
221 470 475 480
222
223 ATC AAG AAG ACA CTA CAA AAA ATT AGC AAC AGT CCA GAG AAG CCT AAA 1782
224 Ile Lys Lys Thr Leu Gln Lys Ile Ser Asn Ser Pro Glu Lys Pro Lys
225 485 490 495 500
226
227 GTG ATT CAA TAGCCCAGGA GCACCTGATT CCTTCTGCC TGCAGGGGGC 1831
228 Val Ile Gln
229
230 TGGGGGGGTG GGGGGCAGTG GATGGTGCCC TATCTGGGTA GAGGTAGTGT GAGTGTGGTG 1891
231
232 TGTGCTGGGG ATGGGCAGCT GCGCCTGCCT GCTCGGCCCC CAGCCCACCC AGCCAAAAAT 1951
233
234 ACAGCTGGGC TGAAACCTGA AAAAAAAAAA AAA 1984
235
236 (2) INFORMATION FOR SEQ ID NO: 2:
237 (i) SEQUENCE CHARACTERISTICS:
238 (A) LENGTH: 503 amino acids
239 (B) TYPE: amino acid
240 (D) TOPOLOGY: linear
241 (ii) MOLECULE TYPE: protein
242
243 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
244
245 Met Thr Leu Gly Ser Pro Arg Lys Gly Leu Leu Met Leu Leu Met Ala
246 1 5 10 15
247
248 Leu Val Thr Gln Gly Asp Pro Val Lys Pro Ser Arg Gly Pro Leu Val
249 20 25 30
250
251 Thr Cys Thr Cys Glu Ser Pro His Cys Lys Gly Pro Thr Cys Arg Gly
252 35 40 45
253
254 Ala Trp Cys Thr Val Val Leu Val Arg Glu Glu Gly Arg His Pro Gln
255 50 55 60
256
257 Glu His Arg Gly Cys Gly Asn Leu His Arg Glu Leu Cys Arg Gly Arg
258 65 70 75 80

```

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/039,177DATE: 11/06/98  
TIME: 17:12:52

INPUT SET: S29620.raw

## \*\*\*\*\* PREVIOUSLY ERRORED SEQUENCES - EDITED \*\*\*\*\*

2767 (2) INFORMATION FOR SEQ ID NO: 29:  
2768 (i) SEQUENCE CHARACTERISTICS:  
2769 (A) LENGTH: 6 amino acids  
2770 (B) TYPE: amino acid  
2771 (D) TOPOLOGY: linear  
2772 (ii) MOLECULE TYPE: peptide  
2773  
2774 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:  
2775  
2776 Gly Thr Lys Arg Tyr Met  
2777 1 5  
2778

---

PAGE: 1

**SEQUENCE VERIFICATION REPORT**  
**PATENT APPLICATION US/09/039,177**

DATE: 11/06/98  
TIME: 17:12:52

***INPUT SET: S29620.raw***

Line

Error

Original Text